Théorie Analytique des Probabilités

Pierre Simon Laplace

Book II §9. pp. 219–228

9. We imagine in an urn r balls marked with the n° 1, r balls marked with n° 2, r balls marked with n° 3, and so on in sequence to the n° n. These balls being well mixed in the urn, one draws them successively; one requires the probability that there will come forth at least one of these balls at the rank[1] indicated by its label[2], or that there will come forth of them at least two, or at least three, etc.

We seek first the probability that there will come forth at least one of them. For this, we will observe that each ball can come forth at its rank only in the first n drawings; one can therefore here set aside the following drawings; now the total number of balls being rn, the number of their combinations n by n, by having regard for the order that they observe among themselves, is, by that which precedes,

$$rn(rn-1)(rn-2)\cdots(rn-n+1);$$

this is therefore the number of all possible cases in the first n drawings.

We consider one of the balls marked with the n° 1, and we suppose that it comes forth at its rank, or the first. The number of combinations of the $rn-1$ other balls taken $n-1$ by $n-1$ will be

$$(rn-1)(rn-2)\cdots(rn-n+1);$$

this is the number of cases relative to the assumption that we just made, and, as

[1] *Translator's note*: This means that a ball marked with 1 will be drawn first, a ball marked with 2 will be drawn second, and so on. In other words, balls will be drawn consecutively by number.

[2] *Translator's note*: The word here is *numéro*, number. However, this refers to the use of a number as a label. In order to distinguish it from *nombre*, number or quantity, I choose to render it as such.

this assumption can be applied to r balls marked with n° 1, one will have

$$r(rn-1)(rn-2)\cdots(rn-n+1)$$

for the number of cases relative to the hypothesis that one of the balls marked with the n° 1 will come forth at its rank. The same result takes place for the hypothesis that any one of the $n-1$ other kinds of balls will come forth at the rank indicated by its label. By adding therefore all the results relative to these diverse hypotheses, one will have

$$(a) \qquad rn(rn-1)(rn-2)\cdots(rn-n+1);$$

for the number of cases in which one ball at least will come forth at its rank, provided however that one removes from them the cases which are repeated.

In order to determine these cases, we consider one of the balls of the n° 1, coming forth first, and one of the balls of the n° 2, coming forth second. This case is contained twice in the preceding number; for it is contained one time in the number of the cases relative to the assumption that one of the balls labeled[3] 1 will come forth at its rank, and a second time in the number of cases relative to the assumption that one of the balls labeled 2 will come forth at its rank; and, as this extends to any two balls coming forth at their rank, one sees that it is necessary to subtract from the number of the cases preceding the number of all the cases in which two balls come forth at their rank.

The number of combinations of two balls of different labels is $\frac{n(n-1)}{1.2}r^2$; for the number of the labels being n, their combinations two by two are in number $\frac{n(n-1)}{1.2}$, and in each of these combinations one can combine the r balls marked with one of the labels with the r balls marked with the other label. The number of combinations of the $rn-2$ balls remaining, taken $n-2$ by $n-2$, by having regard for the order that they observe among themselves, is

$$(rn-2)(rn-3)\cdots(rn-n+1);$$

thus the number of cases relative to the assumption that two balls come forth at their rank is

$$\frac{n(n-1)}{1.2}r^2(rn-2)(rn-3)\cdots(rn-n+1);$$

[3] *Translator's note*: The word is *numérotées*, numbered. I have chosen to render it as such for the same reason as above.

subtracting from it the number (*a*), one will have

$$(a') \qquad \begin{cases} rn(rn-1)(rn-2)\cdots(rn-n+1) \\ \quad -\dfrac{n(n-1)}{1.2} r^2 (rn-2)(rn-3)\cdots(rn-n+1), \end{cases}$$

for the number of all the cases in which one ball at least will come forth at its rank, provided that one subtracts again from this function the repeated cases, and that one adds to them those which are lacking.

These cases are those in which three balls come forth at their rank. By naming k this number, it is repeated three times in the first term of the function (*a*′); for it can result, in this term, from three assumptions of each of the three balls coming forth at its rank. The number k is likewise contained three times in the second term of the function; for it can result from each of the assumptions relative to any two of the three balls coming forth at their rank. Thus, this second term being affected with the $-$ sign, the number k in not found in the function (*a*′); it is necessary therefore to add it to (*a*′) in order that it contain all the cases in which one ball at least comes forth at its rank. The number of combinations of n labels taken three by three is $\frac{n(n-1)(n-2)}{1.2.3}$, and, as one can combine the r balls of one of these labels of each combination with the r balls of the second label and with the r balls of the third label, one will have the total number of combinations in which three balls come forth at their rank, by multiplying $\frac{n(n-1)(n-2)}{1.2.3} r^3$ by $(rn-3)(rn-4)\cdots(rn-n+1)$, a number which expresses that of the combinations of the $rn-3$ balls remaining, taken $n-3$ by $n-3$, by having regard for the order that they observe among themselves. If one adds this product to the function (*a*′), one will have

$$(a'') \qquad \begin{cases} rn(rn-1)(rn-2)\cdots(rn-n+1) \\ \quad -\dfrac{n(n-1)}{1.2} r^2 (rn-2)(rn-3)\cdots(rn-n+1), \\ \quad +\dfrac{n(n-1)(n-2)}{1.2.3} r^3 (rn-3)(rn-4)\cdots(rn-n+1). \end{cases}$$

This function expresses the number of all cases in which one ball at least comes forth at its rank, provided that one subtracts from it again the repeated cases. These cases are those in which four balls come forth at their rank. By applying here the preceding reasonings, one will see that it is necessary again to subtract from the function (*a*″) the term

$$\frac{n(n-1)(n-2)(n-3)}{1.2.3.4}r^4(rn-4)(rn-5)\cdots(rn-n+1).$$

By continuing thus, one will have, for the expression of the cases in which one ball at least comes forth at its rank,

$$(A)\quad\left\{\begin{array}{l} rn(rn-1)(rn-2)\cdots(rn-n+1) \\ \quad -\dfrac{n(n-1)}{1.2}r^2(rn-2)(rn-3)\cdots(rn-n+1) \\ \quad +\dfrac{n(n-1)(n-2)}{1.2.3}r^3(rn-3)(rn-4)\cdots(rn-n+1) \\ \quad -\dfrac{n(n-1)(n-2)(n-3)}{1.2.3.4}r^4(rn-4)(rn-5)\cdots(rn-n+1) \\ \quad +\cdots \end{array}\right.$$

a series being continued as far at it can be. In this function, each combination is not repeated: thus the combination of s balls coming forth at their rank is found here only one time; for this combination is contained s times in the first term of the function, since it can result from each of the s balls coming forth at its rank; it is subtracted $\frac{s(s-1)}{1.2}$ times in the second term, since it can result from two by two combinations of the s balls coming forth at their rank; it is added $\frac{s(s-1)(s-2)}{1.2.3}$ times in the third term, since it can result from the combinations of s letters taken three by three, and so in sequence; it is therefore, in the function (A), contained a number of times equal to

$$s-\frac{s(s-1)}{1.2}+\frac{s(s-1)(s-2)}{1.2.3}-\cdots,$$

and consequently equal to $1-(1-1)^s$, or to unity. By dividing the function (A) by the number $rn(rn-1)(rn-2)\cdots(rn-n+1)$ of all possible cases, one will have, for the expression of the probability that one ball at least will come forth at its rank,

$$(B)\quad\left\{\begin{array}{l} 1-\dfrac{(n-1)r}{1.2(rn-1)}+\dfrac{(n-1)(n-2)r^2}{1.2.3(rn-1)(rn-2)} \\ \quad -\dfrac{(n-1)(n-2)(n-3)r^3}{1.2.3.4(rn-1)(rn-2)(rn-3)}+\cdots \end{array}\right.$$

We seek now the probability that s balls at least will come forth at their rank. The number of cases in which s balls come forth a their rank is, by that which precedes,

$$(b)\quad \frac{n(n-1)(n-2)\cdots(n-s+1)}{1.2.3\ldots s} r^s (rn-s)(rn-s-1)\cdots(rn-n+1),$$

provided that one subtracts from this function the cases which are repeated. These cases are those in which $s+1$ balls come forth at their rank, for they can result, in the function, from $s+1$ balls taken s by s; these cases are therefore repeated $s+1$ times in this function; consequently it is necessary to subtract them s times. Now the number of cases in which $s+1$ balls come forth at their rank is

$$\frac{n(n-1)(n-2)\cdots(n-s)}{1.2.3\ldots(s+1)} r^{s+1} (rn-s-1)(rn-s-2)\cdots(rn-n+1).$$

By multiplying it by s and subtracting it from the function (b), one will have

$$(b')\left\{\begin{aligned} &\frac{n(n-1)(n-2)\cdots(n-s+1)}{1.2.3\ldots s} r^s (rn-s)(rn-s-1)\cdots(rn-n+1) \\ &\qquad \times \left[1 - \frac{s(n-s)r}{(s+1)(rn-s)}\right]. \end{aligned}\right.$$

In this function, many cases are again repeated, namely, those in which $s+2$ balls come forth at their rank; for they result, in the first term, from $s+2$ balls coming forth at their rank and taken s by s; they result, in the second term, from $s+2$ balls coming forth at their rank and taken $s+1$ by $s+1$, and moreover multiplied by the factor s, by which one has multiplied the second term. They are therefore contained in this function the number of times $\frac{(s+2)(s+1)}{1.2} - s(s+2)$; thus it is necessary to multiply by unity, less this number of times, the number of cases in which $s+2$ balls come forth at their rank. This last number is

$$\frac{n(n-1)(n-2)\cdots(n-s-1)}{1.2.3\ldots(s+2)} r^{s+2} (rn-s-2)(rn-s-3)\cdots(rn-n+1);$$

the product in question will be therefore

$$\frac{n(n-1)(n-2)\cdots(n-s-1)}{1.2.3\ldots(s+2)} r^{s+2} (rn-s-2)\cdots(rn-n+1)\frac{s(s+1)}{1.2}.$$

By adding it to the function (b'), one will have

$$(b'')\left\{\begin{array}{l} \dfrac{n(n-1)(n-2)\cdots(n-s+1)}{1.2.3\ldots s} r^s (rn-s)(rn-s-1)\cdots(rn-n+1) \\ \quad \times \left\{\begin{array}{l} 1 - \dfrac{s}{(s+1)} \, \dfrac{(n-s)r}{(rn-s)} \\ \quad + \dfrac{s}{s+2} \, \dfrac{(n-s)(n-s-1)r^2}{1.2(rn-s)(rn-s-1)} \end{array}\right\}. \end{array}\right.$$

This is the number of all possible cases in which s balls come forth at their rank, provided that one subtracts from it again the cases which are repeated. By continuing to reason so, and by dividing the final function by the number of all possible cases, one will have, for the expression of the probability that s balls at least will come forth at their rank,

$$(\mathrm{C})\quad\left\{\begin{array}{l} \dfrac{(n-1)(n-2)\cdots(n-s+1)r^{s-1}}{1.2.3\ldots s(rn-1)(rn-2)\cdots(rn-s+1)} \\ \quad \times \left\{\begin{array}{l} 1 - \dfrac{s}{s+1} \, \dfrac{(n-s)r}{rn-s} + \dfrac{s}{s+2} \, \dfrac{(n-s)(n-s-1)r^2}{1.2.(rn-s)(rn-s-1)} \\ \quad - \dfrac{s}{s+3} \, \dfrac{(n-s)(n-s-1)(n-s-2)r^3}{1.2.3(rn-s)(rn-s-1)(rn-s-2)} + \cdots \end{array}\right\}. \end{array}\right.$$

One will have the probability that none of the balls will come forth at its rank by subtracting the formula (B) from unity, and one will find, for its expression,

$$\frac{(1.2.3\ldots rn) - nr[1.2.3\ldots(rn-1)] + \frac{n(n-1)}{1.2} r^2 [1.2.3\ldots(rn-2)] - \cdots}{1.2.3\ldots rn}.$$

One has, by n° 33[4] of Book I, whatever be i,

$$1.2.3\ldots i = \int x^i dx c^{-x},$$

the integral[5] being taken from x null to x infinity. The preceding expression can therefore be put under this form

[4] pages 128–137.
[5] *Translator's note*: The constant c denotes e, the base of the natural logarithm.

$$\text{(o)} \qquad \frac{\int x^{rn-n}dx(x-r)^n c^{-x}}{\int x^{rn}dx c^{-x}}.$$

We suppose the number rn of balls in the urn very great; then, by applying to the preceding integrals the method of n° 24[6] of Book I, one will find very nearly for the integral of the numerator,

$$\frac{\sqrt{2\pi}X^{rn+2}\left(1-\frac{r}{X}\right)^{n+1}c^{-X}}{\sqrt{n}X^2+n(r-1)(X-r)^2},$$

X being the value of x which renders a maximum the function $x^{rn-n}(x-r)^n c^{-x}$. The equation relative to this maximum gives for X the two values

$$X=\frac{rn+r}{2}\pm\frac{\sqrt{r^2(n-1)^2+4rn}}{2}.$$

One can consider here only the greatest of these values which is, to the quantities nearly of the order $\frac{1}{rn}$, equal to $rn+\frac{n}{n-1}$; then the integral of the numerator of the function (*o*) becomes nearly

$$\frac{\sqrt{2\pi}(rn)^{rn+\frac{1}{2}}c^{-rn}\left(1-\frac{1}{n}\right)^{n+1}\sqrt{r}}{\sqrt{(r-1)(1-\frac{1}{n})^2+1}}.$$

The integral of the denominator of the same function is, by n° 33, quite nearly,

$$\sqrt{2\pi}(rn)^{rn+\frac{1}{2}}c^{-rn};$$

the function (*o*) becomes thus

$$\frac{\left(1-\frac{1}{n}\right)^{n+1}\sqrt{r}}{\sqrt{(r-1)(1-\frac{1}{n})^2+1}}.$$

One can put it under the form

[6] pages 94–96.

$$\frac{\left(1-\frac{1}{n}\right)^{n+1}}{\sqrt{(1-\frac{1}{n})^2+\frac{2}{rn}-\frac{1}{rn^2}}};$$

rn being supposed a very great number, this function is reduced quite nearly to this very simple form

$$\left(\frac{n-1}{n}\right)^n.$$

This is therefore the expression approached more and more by the probability that none of the balls of the urn will come forth at its rank, when there is a great number of balls. The hyperbolic logarithm of this expression being

$$-1-\frac{1}{2n}-\frac{1}{3n^2}-\cdots,$$

one sees that it always goes on increasing in measure as n increases; that it is null, when $n=1$, and that it becomes $\frac{1}{c}$, when n is infinity, c being always the number of which the hyperbolic logarithm is unity.

We imagine now a number i of urns each containing the number n of balls, all of different colors, and that one draws successively all the balls in each urn. One can, by the preceding reasonings, determine the probability that one or more balls of the same color will come forth at the same rank in the i drawings. In reality, we suppose that the ranks of the colors are settled after the complete drawing of the first urn, and we consider first the first color; we suppose that it comes forth the first in the drawings of the $i-1$ other urns. The total number of combinations of the $n-1$ other colors in each urn is, by having regard for their situation among themselves, $1.2.3\ldots(n-1);$ thus the total number of these combinations relative to $i-1$ urns is $[1.2.3\ldots(n-1)]^{i-1};$ this is the number of cases in which the first color is drawn the first altogether from all these urns, and, as there are n colors, one will have

$$n[1.2.3\ldots(n-1)]^{i-1}$$

for the number of cases in which one color at least will arrive at its rank in the drawings from the $i-1$ urns. But there are in this number some repeated cases; thus the case where two colors arrive at their rank in these drawings are contained twice in this number; it is necessary therefore to subtract them from it. The number of these cases is, by that which precedes,

$$\frac{n(n-1)}{1.2}[1.2.3\ldots(n-2)]^{i-1};$$

by subtracting it from the preceding number, one will have the function

$$n[1.2.3\ldots(n-1)]^{i-1} - \frac{n(n-1)}{1.2}[1.2.3\ldots(n-2)]^{i-1}.$$

But this function contains itself some repeated cases. By continuing to exclude from them as one has made above relatively to a single urn, by dividing next the final function by the number of all possible cases, and which is here $(1.2.3\ldots n)^{i-1}$, one will have, for the probability that one of the $n-1$ colors at least will come forth at its rank in the $i-1$ drawings which follow the first,

$$\frac{1}{n^{i-2}} - \frac{1}{1.2[n(n-1)]^{i-2}} + \frac{1}{1.2.3[n(n-1)(n-2)]^{i-2}} - \cdots,$$

an expression in which it is necessary to take as many terms as there are units in n. This expression is therefore the probability that at least one of the colors will come forth at the same rank in the drawings of i urns.